IJSPT

ORIGINAL RESEARCH

# RELIABILITY OF A MOVEMENT QUALITY ASSESSMENT TOOL TO GUIDE EXERCISE PRESCRIPTION (MOVEMENTSCREEN)

**Hunter Bennett, MExSc[1]**
**Kade Davison, PhD[1]**
**John Arnold, PhD[1]**
**Max Martin, BExSc[1]**
**Scott Wood, BExSc[1]**
**Kevin Norton, PhD[1]**

## ABSTRACT

***Background/Purpose:*** Movement quality is commonly assessed to identify movement limitations and guide exercise prescription. Rapid growth in the movement assessment landscape has led to the development and utilization of various movement quality assessments, many without reliability estimates. MovementSCREEN is a novel, tablet-based, video-recorded movement assessment tool, currently without published reliability information. Therefore, the purpose of this study was to determine the intra and inter-rater reliability of the MovementSCREEN, including the impact of rater experience, and provide estimates of measurement error and minimal detectable change.

***Study Design:*** Cross-sectional design; reliability study.

***Methods:*** Thirty healthy young adults (14M:16F, mean age 28.4 yrs, SD 9.1) were video recorded completing the nine MovementSCREEN assessment items on two occasions, two weeks apart. Each individual movement was assessed against objective scoring criteria (component items: yes/no) and using a 100-point sliding scale. To create an overall score for each movement, the scale score is weighted against the objective items to provide a score out of 100. At the completion of all nine individual movements, a mean composite score of movement quality is also established (0-100). The first recording was scored twice by two expert and two novice assessors to investigate inter- and intra-rater reliability. The second recording was scored by one expert assessor to investigate within-subject error. Inter- and intra-rater reliability was calculated using intraclass correlation coefficients (ICCs) and Kappa statistics. The standard error of measurement (SEM), and minimal detectable change ($MDC_{95}$) for the overall score for each movement, and the composite score of movement quality, were calculated.

***Results:*** Intra-rater reliability for the component items ranged from $\kappa$ = 0.619 – 1.000 (substantial to near perfect agreement) and 0.233 – 1.000 (slight to near perfect agreement) for expert and novice assessors, respectively. The ICCs for the overall movement quality scores for each individual movement ranged from 0.707 – 0.952 (fair to high) in expert and 0.502 – 0.958 (poor to high) in novice assessors. Inter-rater agreement for the component items between expert assessors ranged from $\kappa$ = 0.242 - 1.000 (slight to almost perfect agreement), while for novice assessors ranged from 0.103 – 1.000 (less than chance to almost perfect agreement). ICCs for the overall scores for each individual movement from expert and novice assessors ranged from 0.294 – 0.851 (poor to good) and 0.249 – 0.775 (poor to fair), respectively. The SEM for the composite score was 2 points, while the $MDC_{95}$ was 6 points, with an ICC 0.901.

***Conclusions:*** The MovementSCREEN can assess movement quality with fair to high reliability on a test-retest basis when used by experienced assessors, although reliability scores decrease in novice assessors. Comparisons between assessors involve greater error. Therefore, the training of inexperienced assessors is recommended to improve reliability.

***Level of Evidence:*** 2b

***Keywords:*** functional movement screening, movement dysfunction, movement quality, movement system

[1] Alliance for Research in Exercise, Nutrition and Activity, Sansom Institute for Health Research, University of South Australia, Adelaide, Australia.

**Conflict of Interest:** Professor Kevin Norton and Max Martin are directors of Movement Screen Pty Ltd, the company that developed the MovementSCREEN movement assessment tool. Neither was involved in any part of the data management or analysis.

Mr Hunter Bennett, Mr Scott Wood, Dr Kade Davison, and Dr John Arnold declare no conflicts of interest.

**CORRESPONDING AUTHOR**
Hunter Bennett
University of South Australia, GPO Box 2471,
Adelaide, SA, 5001
E-mail: hunter.bennett@mymail.unisa.edu.au
+61 433 377 222

## INTRODUCTION

The assessment of movement quality has become commonplace in both sport and recreational fitness settings, often as tools to predict injury risk.[1,2] While the association between movement quality and injury risk has been inconclusive,[3,4] these assessments can provide coaches, trainers, and rehabilitation practitioners with valuable information regarding areas of muscular weakness, tightness, and movement dysfunction.[1] This information can therefore play an important role in guiding exercise prescription to meet the individual needs of an athlete or client.[1,5]

Increasing interest in assessing movement quality has led to the development and widespread utilization of several movement screening tools in both research and practical settings.[1,2] These tools have largely been developed for use in specific populations with relatively specific objectives, although their overreaching goal is to assess movement quality through the appraisal of an individual's capacity to perform fundamental movements.[1,5,6,7] The inability to complete these movements may be indicative of movement dysfunction, while their successful completion demonstrates a higher level of movement quality.[5,8]

MovementSCREEN is a new electronic-based, video-recorded movement quality assessment tool that assists in gathering information necessary to guide individualized exercise interventions; providing a clear starting point from which an athlete or individual can commence gym-based resistance exercise. MovementSCREEN evaluates the performance of nine fundamental movements. Each individual movement is scored against objective criteria in combination with an overall movement quality indicator to provide an indication of global movement quality. This tool is further stated to assist coaches and practitioners track changes in movement quality that occur in response to individualised exercise interventions. Although its assessment items share many similarities with other movement assessment tools, MovementSCREEN provides a novel, tablet-based method of assessing movement quality that offers simple usability in exercise prescription settings. Additionally, it uses a 100-point scale to provide a composite movement measure: thereby attempting to improve upon a criticism of existing movement quality assessments in their use of basic Likert-type scoring systems (i.e. 0-3), in which a lack of sensitivity has been observed.[9] While the use of these Likert-type scoring systems may have the potential to increase the tools' ease of use, the lack of sensitivity could limit the depth of information gathered, mask potential associations with physical performance and injury risk, and inhibit the ability to track training induced changes in movement quality.[9]

To allow the confident assessment of changes in an individual's movement quality, the tool needs to be reliable. As assessor experience has also been shown to influence the reliability of movement quality assessments,[10] it is important to include raters with different levels of experience when determining reliability. This has implications for increasing the utility of the tool in the field and establishing relevant training of inexperienced assessors. Moreover, while assessing the same movement performance on two separate occasions (via video capture) provides reliability information pertaining to the technical error associated with use of the tool itself, it doesn't provide any information about within-subject error. Within-subject error is likely to be introduced when an individual performs the movement assessment at two different time points, where small variations in movement may be observed.[7] Subsequently, reliability measures that have relevance to clinical practice such as minimum detectable change (MDC95) and standard error measurement (SEM) should be established accounting for within-subject error.

The first aim of this study was to determine the intra and inter-rater reliability of a novel, tablet-based movement quality assessment tool (MovementSCREEN), including estimates of typical measurement error and minimal detectable change. The second aim was to determine the impact that assessor experience has on reliability estimates

## METHODS

### Participants

Participants qualified as apparently healthy in accordance to the Exercise and Sport Science Australia (ESSA) pre-exercise screening tool,[11] were free of musculoskeletal and neurological disease, and were physically able to perform the nine movements within the assessment protocol. A sample size calculation was performed and indicated that with each subject measured two times, a target ICC of 0.9, an ICC of 0.75 or higher to be minimally acceptable, $\alpha = 0.05$ and 80% power, 26 subjects were required.

**Ethical Approval**

This study was approved by the University of South Australia human research ethics committee (0000036268). All participants were informed of the risks and benefits of the investigation prior to signing an institutionally approved consent document to participate in the study. Reporting for this study was conducted in accordance to COSMIN checklist for reliability studies.[12]

**MovementSCREEN Assessment Tool**

Further detail surrounding the application of MovementSCREEN and the rationale of its included movement assessments are provided in supplementary digital content 1. In short, MovementSCREEN was designed to provide coaches, trainers, and exercise professionals a movement quality assessment that indicates an individual's current capacity for gym-based exercise. This information may guide exercise prescription, while also providing a way to quantify changes in movement quality that occur in response to exercise interventions.

The MovementSCREEN protocol involves the assessment of both video-recorded lower-body and upper-body dominant movements, which include: squat, lunge, deadlift with bent over row, single leg squat, overhead reach, thoracic rotation, four-point with opposite arm/leg lift, push up, and active straight leg raise. These nine movements (Figures 1, 2, and 3) are scored individually (Table 1), and they provide an indication of trunk and hip stability, hip, ankle, and thoracic range of motion, and bilateral and unilateral movement control. The movements themselves, and the assessment criteria of those movements (Table 1), provide a means of identifying movement dysfunction. Additionally, as they are movements commonly trained within a gym setting with the intent to improve physical performance and functional capacity, they are relevant to guiding gym-based exercise prescription.[13,14,15] Limitations in movement quality have been suggested to present through an inability to complete a given movement in accordance to its set scoring criteria 1 (Table 1). Dependant on that scoring criteria, these limitations will be indicative of the site and severity of potential movement dysfunction. This information can then provide a basis for further assessment if required, guiding subsequent exercise prescription.[13,14,15]

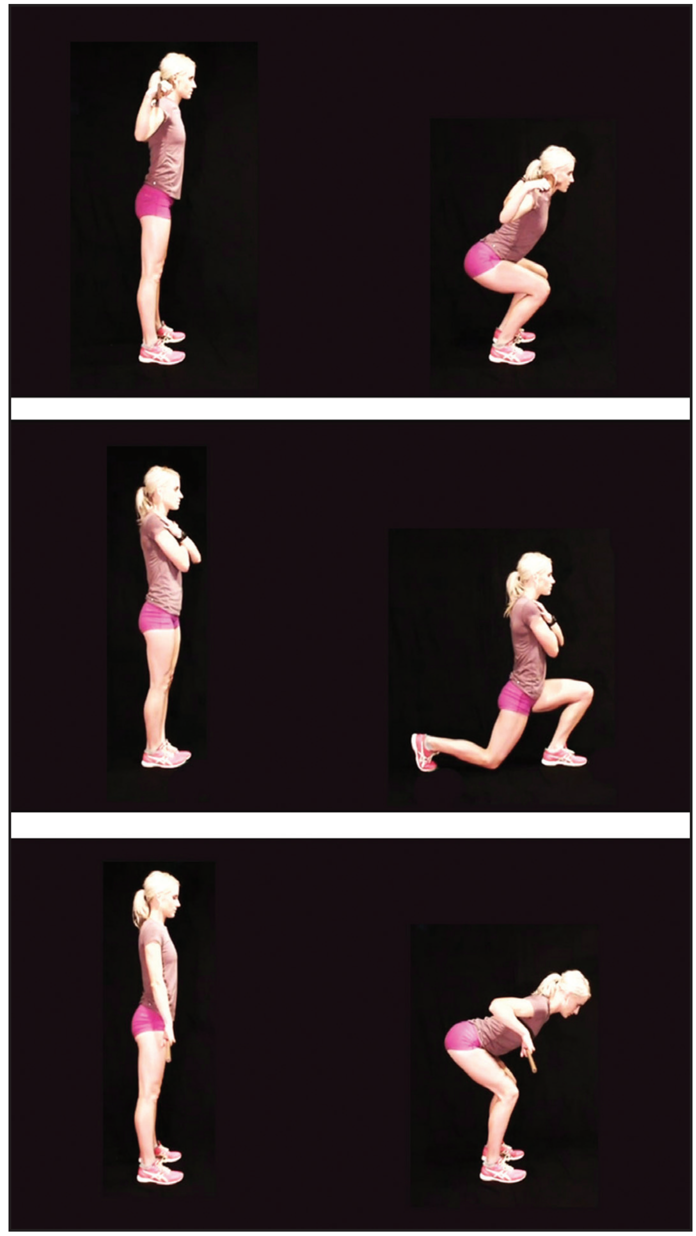

**Figure 1.** *Side view of start and finish position of the 1) Squat, 2) the Lunge, and the 3) Deadlift with bent over row.*

Within this assessment, each movement contains specific criteria that are presented as component items and are scored as a 'yes' or 'no' response

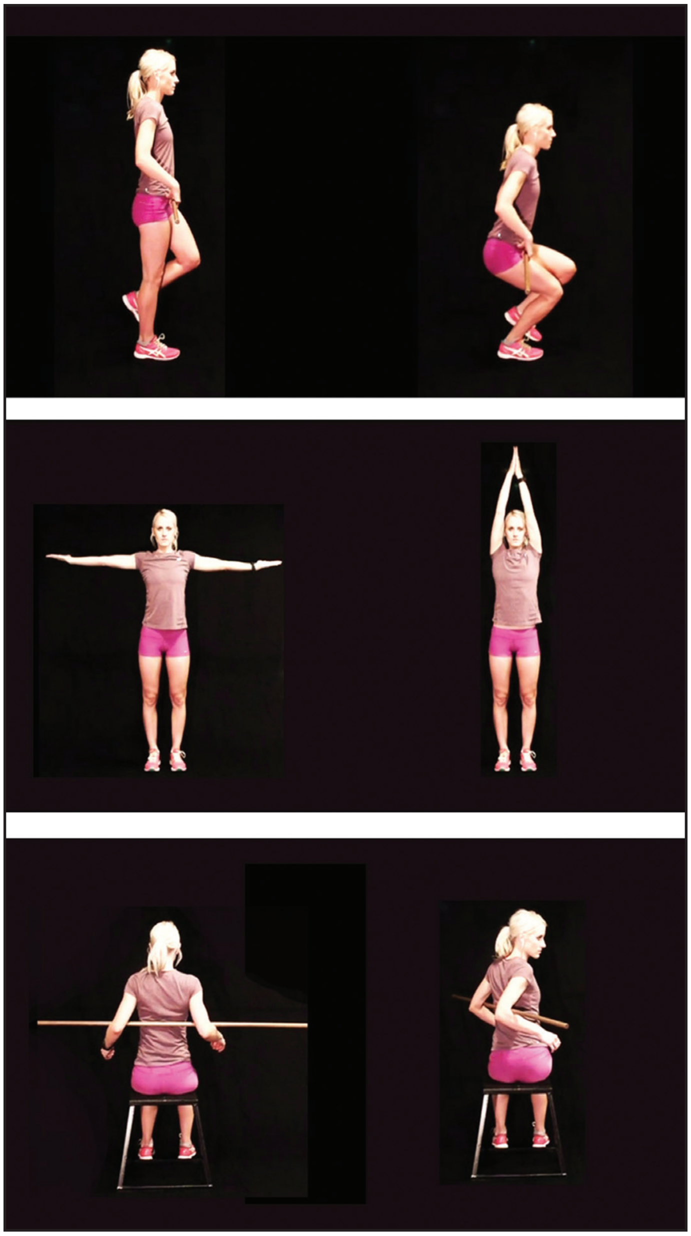

**Figure 2.** *Side view of start and finish position of the 1) Single leg squat, 2) the Overhead reach, and the 3) Thoracic rotation.*

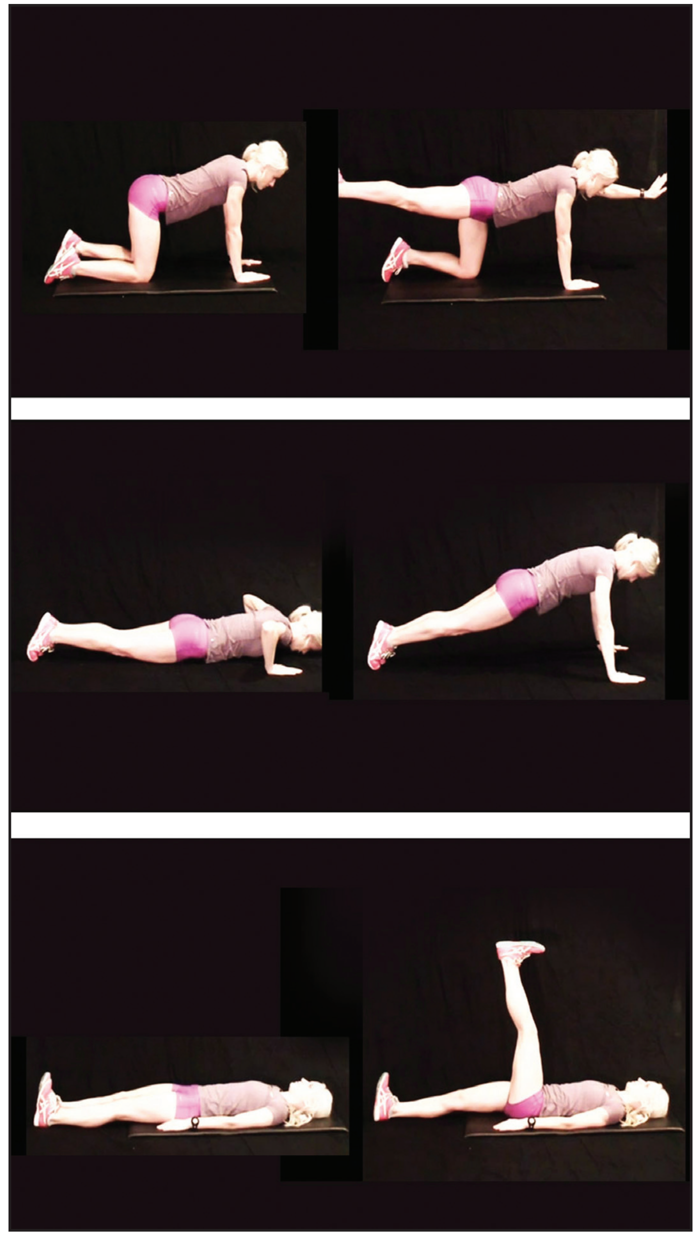

**Figure 3.** *Side view of start and finish position of the 1) 4-Point with opposite arm/leg lift, 2) the push up, and the 3) active straight leg raise.*

(Table 1). The component items relate to important aspects of each movement and are based on elements of control required for safe and effective movement. The quality of each individual movement is also scored using a 100-point sliding scale with associated cues (with a score of 100 being indicative of perfect movement quality). To create a final movement quality score for each movement, the subjective score is weighted against the sum of the component items to provide an overall score out of 100 (100 being the highest achievable score). For unilateral movements, each side is scored separately, and a mean score of the two sides is provided for the overall movement quality score for that specific movement. At the completion of all nine movements, a mean composite score is calculated from the overall scores of each induvial movement to provide a global measure of movement quality (0-100).

### *Protocol*

Participants were video recorded completing the MovementSCREEN assessment protocol twice,

**Table 1.** *MovementSCREEN assessment protocol.*

| Movement | Assessment Set Up | Movement Cues | Objective Scoring Criteria (Yes/No) |
|---|---|---|---|
| **Squat** | Client to place feet shoulder width apart with ~10 degrees of external rotation.<br>Dowel to be placed and held behind the neck. | Squat down as low as you comfortably can as if you're sitting back into a chair. Aim for thighs to reach horizontal. Ensure heels remain on the ground. | • Neutral lumbar position (no hyperextension of flexion)<br>• Trunk parallel to shins<br>• Thighs reached horizontal plane<br>• Dowel remains over foot<br>• Vertical pelvic displacement<br>• Knees in line with feet |
| **Lunge** | At the bottom position of the lunge, both knees should be flexed to 90 degrees, with the grounded knee directly below the hip (mark front and back foot in this position).<br>The movement is performed with feet hip-width apart and facing forward with hands across chest. | Take a step forward so that your front foot's toes land as close to the far marker as possible. Drop into a lunge as far down as you feel comfortable.<br>The movement is scored up to the point that the client reaches the bottom of the lunge. | • Neutral lumbar position (no hyperextension or flexion)<br>• Trunk parallel to front shin<br>• Lead thigh reaches horizontal plane<br>• Horizontal pelvis and shoulders<br>• Knees in line with feet<br>• Balance maintained |
| **Deadlift with bent over row** | Client to place feet shoulder width apart with ~10 degrees external rotation. Hold onto the dowel with the arms and hands hanging by the side of the body. | Bend down by bending ('hinging') at the hips, while keeping the spine/back 'neutral'.<br>Bend down as far as comfortable, aiming for the trunk to be parallel to the ground. Keep heels on ground. Once bent over perform a 'rowing' movement, bringing the dowel to your chest. | • Trunk to parallel<br>• Trunk/spine remains in neutral alignment, no change to lumbar or thoracic curvature<br>• Symmetrical weight loading (movement path)<br>• Knees stay in line with feet<br>• Hands/dowel to chest height<br>• Horizontal body and dowel alignment |
| **Single leg Squat** | The client stands with their feet together and hold a dowel in front of the hips and with thumbs at the level of the anterior superior iliac spines [ASIS] to facilitate viewing of pelvic alignment. | Lift one foot off ground by bending at the knee and keeping thighs parallel. Squat down on one leg by sitting back as far as comfortable, aiming for a 90 degree angle at the knee. Keep heel on the ground. | • Maintains neutral spine alignment<br>• Trunk parallel to grounded leg shin<br>• Reaches 90 degrees at knee<br>• Dowel remains on frontal plane and horizontal<br>• Hip/knee/foot alignment is maintained<br>• Balance is maintained |
| **Overhead Reach** | The client stands with feet hip-width apart. Arms should be extended to the sides with palms facing upwards. | Elbows remain locked in extension throughout the movement's duration. Abduct the arms as far as possible, aiming for the palms to touch overhead. If observations are different for each shoulder, score the movement according to the 'worse' side. | • Palms touch overhead<br>• No obvious hitching of shoulders<br>• Arms are on frontal plane in line with ears<br>• Lumbar spine remains neutral<br>• Side plumb alignment remains neutral |
| **Thoracic rotation** | Instruct the client to sit on a bench/stool that allows the hips to be flexed at 90 degrees, with an upright and neutral spine. Place the dowel behind the back, which is then held in the crooks of the elbows, with forearms facing forwards | Smoothly rotate to one side as far as possible, and then the other, while maintaining a neutral spine.<br>The pelvis or legs should not move to achieve a greater rotation. | • Maintains neutral vertical alignment |

| **Table 1.** *MovementSCREEN assessment protocol. (continued)* | | | |
|---|---|---|---|
| **Movement** | **Assessment Set Up** | **Movement Cues** | **Objective Scoring Criteria (Yes/No)** |
| **4 point arm/leg lift** | The individual kneels on the ground/mat on 'all fours'. Wrists are kept under the shoulders, and knees under hips, with a neutral spinal position. | Lift one knee off the ground by extending at the hip and knee while lifting the opposite hand and flexing at the shoulder and elbow. Hold the position for one second and return to starting position. Repeat with the other leg/hand combination. The side being scored is the grounded arm side. | • Maintains balance while one leg/one hand on ground<br>• Maintains neutral spine throughout<br>• Maintains level rib cage<br>• Maintains level pelvis<br>• Grounded scapula remains stable on ribcage |
| **Push Up** | The individual starts lying on their front on the ground, with hands flat on floor and level with armpits, sitting slightly wider than shoulders. Toes are on floor and flexed. The push up will be performed by raising up from that set position and lowering back down with control. | Perform a push up by pushing up from the floor and lowering back down with control. Aim to keep the body in a straight line from ankle, hips, to shoulders, so the hips don't 'fall behind' (prevent the hips from sagging). If a push up is unable to be completed, try a push up from the knees on the floor. If this is still not possible, use a standing position with hands on a secure surface at a height of the anterior superior iliac spine [ASIS]. | • Maintains neutral spinal alignment<br>• Shoulders neutral and not hitched<br>• Hips stable and aligned with trunk<br>• Head neutral in line with shoulders<br>• No obvious scapular winging |
| **Active Straight Leg Raise** | The client starts by lying on the floor on their back with legs together and arms by the side. | Lift one leg off the ground by flexing at the hip while keeping the knee extended. Ensure that the grounded leg remains fully extended and does not lift from the floor. Observe the height that the leg is raised and replicate/score this on the slider below the 'drawn' figure. The distance scored is the % change in the position of the dowel relative to 100% distance achieved [where the lifted leg is at 90 degrees to the ground]. | N/A |

two weeks apart. The participants performed the MovementSCREEN protocol in the sequential order outlined in Table 1 and illustrated in Figures 1-3. Participants were given specific instructions on how to perform each movement, while also observing a filmed demonstration of the movements performed with optimal technique. A short warmup was performed prior to the assessment which included a five-minute bout of jogging, followed by some bodyweight exercises (walking lunges with arms overhead, leg swings, and overhead reaches). Feedback during the movement tasks was prohibited. The entire assessment, including the warmup, took approximately 30 minutes per participant.

Each testing session was video recorded using two Apple iPads (30 frames per second, 1080p) mounted on tripods and positioned four meters from the participant. Cameras were positioned orthogonal to each other, where one camera recorded the sagittal plane of movement and one the frontal plane from the anterior aspect. Each participant performed this protocol on two individual occasions separated by 14 days.

Two expert and two novice assessors assessed the first video recording twice, 14 days apart to investigate intra- and inter- rater reliability. The expert assessors were university-qualified exercise science graduates, each with over five years of experience working clinically as exercise physiologists, and strength and conditioning specialists. The novice assessors were current clinical exercise physiology students in their final year of study, with knowledge of exercise prescription practices, but relatively little practical experience assessing and prescribing exercise. The video footage of the second session was also assessed by one of the expert assessors to investigate the reliability estimates for within-subject error. In this manner, the assessment of the two time points accounted for both the technical error of assessment and the performer's variability in movement from one test to another, which cannot be considered when only assessing video footage from a single time point. All assessments for all assessors were performed under the same conditions, with the filmed video footage being re-watched on an 18-inch computer monitor independently.

**Statistical Analysis**

Descriptive statistics were calculated for both the overall and component scores for each tester and session. All descriptive data are presented as mean and standard deviations, where appropriate. Paired t-tests were performed on both the movement and composite scores for one expert and one novice assessor to assess for systematic error. The intra- and inter- rater reliability of each individual assessment item within each movement (binary data: Yes/No) were determined with the Kappa statistic. Intraclass correlation coefficients (ICC) were used to measure the agreement between the two groups for the overall scores. Intra- and inter-rater reliability was calculated with a two-way mixed ICC for the overall scores between expert and novice raters separately. Within-subject error was established using ICC to measure the agreement between testing sessions one and two, as scored by an expert assessor.

Response stability of the overall scores of each individual movement, and the composite scores was calculated using the standard error of the measurement (SEM) at the 95% level of confidence.[16] The minimal detectable change (MDC95) values at the 95% level of confidence were calculated to determine the lowest level of change that can be considered 'true' change and not likely due to measurement error 19. ICC's were interpreted according to the following criteria: high (0.90–0.99); good (0.80–0.89); fair (0.70–0.79) and poor (0.00–0.69).[17] Kappa statistics were interpreted according to Landis and Koch:[18] slight agreement (0.01-0.20), fair agreement (0.21- 0.40), moderate agreement (0.41-0.60), substantial agreement (0.61-0.80), and almost perfect agreement (0.81-1.00).

Data were analysed using the statistical package SPSS version 24.0 for Windows, PC (IBM, Chicago, IL). Alpha was set at the 0.05 level.

**Results**

Thirty apparently healthy adults (m = 14, f = 16; mean age 28.4 years, SD 9.1; mean height 171.3 cm, SD 9.4; mean weight 70.5 kg, SD 12.7) participated in this study. All 30 participants recruited into the study that met the inclusion criteria completed data collection, with no dropouts.

Table 2 provides the mean scores for both novice and expert assessors for each movement, and demonstrates the intra- rater reliability for both the component items and movement quality scores. Kappa scores for the component items within the expert raters ranged from 0.619 – 1.000, suggesting substantial to near perfect agreement, while the novice raters demonstrated slight to near perfect agreement (0.233 – 1.000). ICCs for the final movement quality score demonstrated fair to high intra-rater reliability in expert assessors (0.707 – 0.952), and poor to high intra-rater reliability (0.502 – 0.958) in novice assessors. Paired t-test across all movement and composite scores demonstrated non-uniform differences across both expert ($p = .022 - .729$) and novice assessors ($p = .039 - .998$), indicating no systematic difference in scores between the first and second assessments.

Table 3 outlines inter- rater agreement between expert and novice assessors. Component items comparisons between expert assessors showed Kappa ranged from 0.242 - 1.000, suggesting slight to almost

**Table 2.** *MovementSCREEN movement scores (mean and standard deviation), and intra-rater reliability for component items and movement quality scores, for novice and expert assessors.*

| | Expert 1 | Expert 2 | Novice 1 | Novice 2 |
|---|---|---|---|---|
| **Squat** | | | | |
| Mean (Standard deviation) | 52 (18) | 44 (15) | 50 (14) | 50 (11) |
| Component items (Kappa) | 0.754 – 1.000 | 0.701 – 0.927 | 0.342 – 1.000 | 0.585 – 1.000 |
| Movement Quality (ICC) | 0.887 (0.779 – 0.945) | 0.835 (0.556 – 0.930) | 0.712 (0.481 – 0.851) | 0.704 (0.467 – 0.847) |
| **Lunge** | | | | |
| Mean (Standard deviation) | 68 (11) | 58 (12) | 63 (11) | 60 (5) |
| Component items (Kappa) | 0.712 – 1.000 | 0.718 – 1.000 | 0.412 – 1.000 | 0.348 – 1.000 |
| Movement Quality (ICC) | 0.715 (0.481 – 0.854) | 0.738 (0.520 – 0.866) | 0.649 (0.383 – 0.816) | 0.652 (0.172 – 0.850) |
| **Deadlift with bent over row** | | | | |
| Mean (Standard deviation) | 52 (10) | 44 (10) | 56 (14) | 53 (7) |
| Component items (Kappa) | 0.762 – 1.000 | 0.737 – 1.000 | 0.526 – 1.000 | 0.526 – 1.000 |
| Movement Quality (ICC) | 0.851 (0.713 – 0.926) | 0.763 (0.479 – 0.890) | 0.803 (0.626 – 0.901) | 0.708 (0.470 – 0.849) |
| **Single leg squat** | | | | |
| Mean (Standard deviation) | 43 (10) | 30 (10) | 42 (8) | 44 (9) |
| Component items (Kappa) | 0.651 – 1.000 | 0.619 – 1.000 | 0.359 – 1.000 | 0.361 – 1.000 |
| Movement Quality (ICC) | 0.902 (0.805 – 0.952) | 0.833 (0.186 – 0.947) | 0.706 (0.466 – 0.849) | 0.769 (0.571 – 0.882) |
| **Overhead reach** | | | | |
| Mean (Standard deviation) | 86 (12) | 56 (23) | 83 (8) | 84 (9) |
| Component items (Kappa) | 1.000 | 0.734 – 1.000 | 0.652 – 1.000 | 0.609 – 1.000 |
| Movement Quality (ICC) | 0.875 (0.732 – 0.941) | 0.884 (0.771 – 0.943) | 0.502 (0.163 – 0.748) | 0.510 (0.119 – 0.730) |
| **Thoracic rotation** | | | | |
| Mean (Standard deviation) | 65 (15) | 43 (15) | 59 (12) | 48 (15) |
| Component items (Kappa) | 1.000 | 1.000 | 1.000 | 1.000 |
| Movement Quality (ICC) | 0.868 (0.732 – 0.936) | 0.936 (0.866 – 0.969) | 0.775 (0.580 – 0.886) | 0.900 (0.801 – 0.951) |
| **4 point opposite arm leg lift** | | | | |
| Mean (Standard deviation) | 57 (9) | 63 (14) | 62 (11) | 69 (8) |
| Component items (Kappa) | 0.651 – 1.000 | 0.772 – 1.000 | 0.233 – 1.000 | 0.430 – 1.000 |
| Movement Quality (ICC) | 0.735 (0.518 – 0.864) | 0.707 (0.474 – 0.848) | 0.533 (0.226 – 0.746) | 0.697 (0.448 – 0.844) |
| **Push up** | | | | |
| Mean (Standard deviation) | 66 (21) | 55 (22) | 61 (20) | 72 (18) |
| Component items (Kappa) | 0.652 – 1.000 | 0.652 – 1.000 | 0.423 – 1.000 | 0.526 – 1.000 |
| Movement Quality (ICC) | 0.928 (0.855 – 0.965) | 0.887 (0.778 – 0.944) | 0.882 (0.766 – 0.942) | 0.885 (0.773 – 0.944) |
| **Active straight leg raise** | | | | |
| Mean (Standard deviation) | 52 (19) | 45 (17) | 55 (19) | 50 (19) |
| Movement Quality (ICC) | 0.910 (0.813 – 0.957) | 0.916 (0.824 – 0.960) | 0.904 (0.808 – 0.953) | 0.827 (0.662 – 0.915) |
| **Composite Score** | | | | |
| Mean (Standard deviation) | 60 (7) | 49 (7) | 59 (7) | 59. (5) |
| Movement Quality (ICC) | 0.952 (0.893 – 0.978) | 0.862 (0.625 – 0.942) | 0.958 (0.914 – 0.980) | 0.875 (0.755 – 0.938) |

perfect agreement. Novice agreement ranged from 0.103 – 1.000, suggesting less than chance to almost perfect agreement. ICC for the final movement quality scores in expert (0.294 – 0.851) and novice (0.249 – 0.775) assessors demonstrated poor to good and poor to fair reliability, respectively.

Within-subject error between testing sessions one and two for the final movement quality scores varied from poor to good (ICC 0.631 – 0.919). The SEM ranged from 10 to 18, while MDC95 ranged from 14 to 26 (Table 4). The ICC for the composite score of global movement quality was 0.901, while SEM was 2 points, and MDC95 was 6 points.

## DISCUSSION

The assessment of movement quality has the capacity to provide coaches and exercise professionals with valuable information that can advise the development and application of exercise interventions.[1] There has been a rapid increase the number of assessment tools used in practice, although many do not have evidence of their reliability. MovementSCREEN is a novel tablet-based movement assessment tool that appears to have practical merit, but no published reliability estimates to support its use. Therefore, the aims of this study were to determine the intra and inter-rater reliability of MovementSCREEN, including estimates of typical

**Table 3.** *Inter-rater reliability for component items and movement quality scores in novice and expert assessors.*

| | Expert | Novice |
|---|---|---|
| **Squat** | | |
| Component items (Kappa) | 0.658 – 1.000 | 0.366 – 1.000 |
| Movement Quality (ICC) | 0.741 (0.355 – 0.889) | 0.660 (0.395 – 0.823) |
| **Lunge** | | |
| Component items (Kappa) | 0.351 – 1.000 | 0.143 – 1.000 |
| Movement Quality (ICC) | 0.485 (0.038 – 0.767) | 0.400 (0.063 – 0.658) |
| **Deadlift with bent over row** | | |
| Component items (Kappa) | 0.534 – 1.000 | 0.189 – 1.000 |
| Movement Quality (ICC) | 0.463 (0.28 – 0.730) | 0.384 (0.046 – 0.647) |
| **Single leg squat** | | |
| Component items (Kappa) | 0.242 – 1.000 | 0.103 – 1.000 |
| Movement Quality (ICC) | 0.474 (-0.097 – 0.785) | 0.433 (0.107 – 0.679) |
| **Overhead reach** | | |
| Component items (Kappa) | 0.595 – 1.000 | 0.268 – 1.000 |
| Movement Quality (ICC) | 0.294 (-0.104 – 0.628) | 0.249 (-0.122 – 0.557) |
| **Thoracic rotation** | | |
| Component items (Kappa) | 1.000 | 1.000 |
| Movement Angle (ICC) | 0.407 (0.071 – 0.770) | 0.579 (-0.071 – 0.843) |
| Movement Quality (ICC) | 0.511 (-0.92 – 0.826) | 0.581 – (0.051 – 0.838) |
| **4 point arm and leg lift** | | |
| Component items (Kappa) | 0.322 – 1.000 | 0.143 – 1.000 |
| Movement Quality (ICC) | 0.499 (0.186 – 0.723) | 0.260 (-0.055 – 0.545) |
| **Push Up** | | |
| Component items (Kappa) | 0.449 – 1.000 | 0.294 – 1.000 |
| Movement Quality (ICC) | 0.726 (0.266 – 0.887) | 0.673 (0.168 – 0.864) |
| **Active straight leg raise** | | |
| Movement Angle (ICC) | 0.857 (0.495 – 0.946) | 0.840 (0.401 – 0.942) |
| Movement Quality (ICC) | 0.851 (0.506 – 0.943) | 0.775 (0.561 – 0.889) |
| **Composite Score** | | |
| Movement Quality (ICC) | 0.512 (00.29 – 0.847) | 0.693 (0.446 – 0.842) |

measurement error and minimal detectable change, and to determine the impact of assessor experience on those reliability estimates. Data collected as part of this study demonstrated that MovementSCREEN can assess movement quality with fair to high reliability on a test-retest basis when used by experienced assessors, although reliability scores decrease in novice assessors. Subsequently, standardized training looks to be necessary to improve reliability in inexperienced assessors. Moreover, the reliability estimates provided can determine whether 'true' changes in movement quality have occurred, and are essential to inform the interpretation of assessment results in the field and future research studies.

Kappa statistics demonstrated substantial to near-perfect agreement on a test-retest basis for the component items within expert assessors, while ICCs for the movement quality scores were fair to high in the same group. Agreement was generally lower in novice assessors for nearly all movements (Table 3). This information alone suggests there is likely to be a learning effect associated with the assessment of movement quality, with experience assessing movement leading to more consistent scoring.[10] Interestingly, the range of scores allocated for each movement were greater in expert assessors than their novice counterparts when viewing the same video. This suggests that with assessment

| Table 4. *Within-subject error of the movement quality assessment, as assessed by a single expert assessor.* | | | | | |
|---|---|---|---|---|---|
| | ICC | SEM | $SEM_{95}$ | MDC | $MDC_{95}$ |
| **Squat** | 0.725 (0.496 – 0.860) | 9 | 18 | 13 | 26 |
| **Lunge** | 0.777 (0.580 – 0.889) | 5 | 10 | 7 | 14 |
| **DL with bent over row** | 0.631 (0.631 – 0.808) | 6 | 12 | 9 | 18 |
| **Single leg squat** | 0.694 (0.441 – 0.844) | 5 | 11 | 8 | 16 |
| **Overhead reach** | 0.839 (0.629 – 0.928) | 5 | 10 | 7 | 14 |
| **Thoracic rotation** | 0.763 (0.554 – 0.882) | 7 | 14 | 10 | 20 |
| **4-point arm and leg lift** | 0.646 (0.378 – 0.816) | 5 | 11 | 8 | 16 |
| **Push up** | 0.919 (0.837 – 0.961) | 6 | 12 | 8 | 17 |
| **Active straight leg raise** | 0.846 (0.558 – 0.937) | 7 | 15 | 10 | 21 |
| **Composite score** | 0.901 (0.802 – 0.952) | 2 | 4 | 3 | 6 |

experience, there is likely to be an increased exposure to large variations in movement quality. This exposure may result in a greater learned 'spectrum' of movement quality, causing greater discrimination during observation and explaining the increased confidence in expert assessors to score an individual either lower or higher than novice assessors.

The reliability estimates for the MovementSCREEN are comparable to other pre-existing movement quality assessment tools that have been evaluated in the literature.[1] The current investigation demonstrated that the inter-rater reliability for the individual component items demonstrated fair to almost perfect agreement in expert assessors, and poor to almost perfect agreement in novice assessors. Comparatively, the movement quality scores demonstrated poor inter-rater reliability in both expert and novice assessors. This suggests that although the 100-point movement quality score may offer a way to capture smaller variations in movement quality, it might be too subjective for the system to be used interchangeably between assessors without appropriate assessment standardization and education, irrespective of their experience. It is possible that coordinated and standardized training may assist in this regard to improve the utility of the system among assessors.[19] The objectivity associated with the component items appears to improve reliability between both novice and expert assessors. As the information gathered from these specific assessment items is arguably more valuable in terms of exercise prescription guidance, this will likely have positive implications in practical settings.

Although ICC's and Kappa statistics provide a good indication of the tool's reliability, SEM and MDC95 may be more useful in practical settings.[20] Additionally, as the SEM and MDC95 scores within this study have been established from two separate testing occasions, they also account for both within-subject error and technical error. The MDC95 of the overall movement quality scores of each individual movement varied between 14 and 26 points (Table 4). While this variability in some cases may appear quite large, at its highest it represents a 26% change in the movement quality score to confidently suggest that a 'true' change has occurred within an individual movement. This is comparable to observing a one-point change in a given movement within the FMS™, which utilises a four-point ordinal scoring scale.[5] The MDC95 for the composite score was substantially lower at six points, suggesting any changes beyond this would represent a 'true' change in global and whole-body movement quality. It is important to note that while the MDC95 value for the composite score was markedly less than those established for the individual movements, there is explanation for this. As the composite score is derived from a mean of the individual movement scores, it is likely to smooth out the variances seen between those scores, resulting in tighter reliability measures. Therefore, taking into consideration both the composite score and overall score for each individual movement is integral when assessing and tracking changes in movement quality.

The study has several identified strengths. Reliability measures were established in both expert and novice assessors to determine how rater experience affects assessment reliability. Both the SEM and MDC95 were established using two individual testing sessions to account for both the technical error of assessment and within-subject error. This better represents how the tool is used in practical settings, and may provide useful information surrounding the interpretation of any changes in movement quality observed from one test to another.

There are also limitations that should be considered when deciphering the results of this study. The participant group consisted of apparently healthy young adults. While variations in movement quality were observed, these were likely to be less than those seen in clinical practice. As such, the reliability results described may not be equivalent for higher level athletic or clinical populations, with further reliability studies required in this area. Nonetheless, it was first necessary to determine reliability in a healthy population before progressing to more specific groups. It is also important to note that while assessing video recorded footage of the MovementSCREEN offers a convenient means to establish inter- and intra rater reliability, it does slightly restrict the overall visual information provided to the assessor when compared to real time assessment scenarios. In real time assessment, the assessor has the capacity to move around the subject as they perform the movement to gather further information if required. Therefore, the reliability measures described may not necessarily depict those obtained in real time assessments. Finally, while the participants did receive thorough instructions and demonstrations surrounding the correct performance of the individual movement assessments, they did not receive any feedback during the movement. As both feedback and knowledge of the grading criteria have been shown to influence movement assessment outcomes,[21] this study implemented the protocol according to intended use in the field in individuals naïve to the criteria. The impact that different levels of cueing and prior knowledge has on reliability and movement quality with this tool would require further investigation.

## CONCLUSIONS

The MovementSCREEN was developed to meet the needs of coaches, exercise, and rehabilitation professionals working within gym-based exercise prescription environments. The results from this analysis suggest that the tool can assess movement quality with fair to high reliability on a test-retest basis when administered by experienced practitioners. Fair to almost perfect inter-rater agreement was observed for the component items of the assessment tool, although the inter-rater reliability for the subjective movement quality score was poor. This suggests that the scores from the tool may not be reliable enough for confident application between different practitioners without standardized training surrounding its use. The $MDC_{95}$ for the composite score of global movement quality is approximately six points, while it varied between 14 and 26 points for the component assessment items. This information is integral to any future research examining the capacity of MovementSCREEN to identify changes in movement quality that occur over time or after interventions. Further research is required to validate its capacity to guide exercise prescription and track associated changes in movement quality.

## REFERENCES


1. Bennett H, Davison K, Arnold J, et al. Multicomponent musculoskeletal movement assessment tools: A systematic review and critical appraisal of their development and applicability to professional practice. *J Strength Cond Res.* 2017;31: 2903-2919.
2. McCunn R, aus der Fünten K, Fullagar H, et al. Reliability and association with injury of movement screens: A critical review. *Sports Med.* 2016;46: 763-781.
3. Moran R, Schneiders A, Mason J, et al. Do functional movement screen (FMS) composite scores predict subsequent injury? A systematic review with meta-analysis. *Br J Sports Med.* 2017;51:1661-1669.
4. Newton F, McCall A, Ryan D, et al. Functional movement screen (FMS™) score does not predict injury in English Premier League youth academy football players. *Sci Med Footbal.* 2017;1: 1-5.
5. Cook G, Burton L, Hoogenboom B, et al. Functional movement screening: the use of fundamental movements as an assessment of function-part 1. *Int J Sports Phys Ther.* 2014a;9:396-409.
6. Kritz M, Cronin J, Hume P. The bodyweight squat: A movement screen for the squat pattern. *Strength Cond J.* 2009a;31(1):76-85.
7. McKeown I, Taylor-McKeown K, Woods C, et al. Athletic ability assessment: a movement assessment protocol for athletes. *Int J Sports Phys Ther.* 2014;9:862-873.

8. Kritz M. *Development, reliability and effectiveness of the movement competency screen (MCS)*. Doctoral dissertation, Auckland University of Technology. 2012.
9. Frost D, Beach T, Campbell T, et al. An appraisal of the functional movement screen™ grading criteria–Is the composite score sensitive to risky movement behavior? *Phys Ther Sport.* 2015;16:324-330.
10. Norton K, Norton L. *Pre-exercise screening, guide to the Australian adult pre-exercise screening system.* Exercise and Sports Science Australia. 2011.
11. Gulgin H, Hoogenboom B. The functional movement screening (FMS)™: An inter-rater reliability study between raters of varied experience. *Int J Sports Phys Ther.* 2014;9:14-20.
12. Mokkink L, Terwee C, Patrick D, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res.* 2010;19:539-549.
13. Cook G. *Movement: Functional movement systems: Screening, assessment, corrective strategies*. On Target Publications. 2010.
14. Haff G, Triplett N. *Essentials of strength training and conditioning 4th edition*. Human kinetics. 2015.
15. American College of Sports Medicine (ACSM). *ACSM's guidelines for exercise testing and prescription 9th Ed*. Lippincott Williams & Wilkins. 2013.
16. Weir J. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res.* 2005;19:231-240.
17. Onate J, Dewey T, Kollock R, et al. Real-time intersession and interrater reliability of the functional movement screen. *J Strength Cond Res.* 2012;26:408-415.
18. Landis J, Koch G. The measurement of observer agreement for categorical data. *Biometrics.* 1977;1:159-174.
19. Starkes J, Ericsson K. *Expert performance in sports: Advances in research on sport expertise*. Human Kinetics. 2003.
20. Haley S, Fragala-Pinkham M. Interpreting change scores of tests and measures used in physical therapy. *Phys Ther.* 2006;86:735-743.
21. Frost D, Beach T, Callaghan J, et al. FMS scores change with performers' knowledge of the grading criteria—are general whole-body movement screens capturing "dysfunction"?' *J Strength Cond Res*. 2015;29:3037-3044.